

**Recommendation:** Major clarification and revision are required

**General remarks:**

In this paper, an Ensemble Transformation Adjoint (ETA) method is proposed to improve the previous ET method in computational efficiency, and numerical experiments are performed to compare ETA with ET. The idea is interesting but not clearly explained, and the technique is new but not cleanly presented. In particular, some important assumptions (or limitations) used implicitly in the derivations are not stated explicitly. The derivations are presented carelessly with errors (typos and/or unexplained omissions), which caused extra efforts to this review (to “decode” the errors). My specific comments and suggestions are given below:

**Specific comments and suggestions:**

1. Since ETA is based on ET of Bishop and Toth (1999), it's better to use the same symbols (such as, with bold upper and lower cases for matrices and vectors, respectively).
2. Descriptions of the derivations are often inaccurate or sloppy. For example, I can list the following problems just from the last paragraph (lines 157-163) on p7:
  - (i) “perturbations” should be “perturbation matrix” in three places on lines 157-158 with “denote” changed to “denotes” on line 157.
  - (ii) For the statement on lines 159-160, it is unclear about the transformation (to what) until (3) is presented on line 163. Here, I suggest to revise “transformation of the ensemble perturbation  $X_e(a)$ ” to “transformation of  $X_e(a)$  to  $Z(a)$ ”.
  - (iii) The statement on lines 160-162 is inaccurate or fuzzy about the uniqueness condition cited from Anderson (1997), and this cited paper is not listed in the reference.

Problems like the above examples are also seen in other places of the paper.

2. In (9) on p11, it is assumed implicitly that each targeting observation (used by the supposedly “optimal” analysis) reduces only the analysis error variance of merely the observed variable at the targeting-observation location. This is a very crude assumption. This assumption not only neglects the true effects of each targeting observation on analysis correlations (between different spatial points and between different variables) but also further implies that the analysis error reductions from different targeting observations are independent when these observations are used in combinations by the “optimal” analysis. This assumption and related limitations should be stated and explained in the paper.

3.  $V$  is neither clearly defined in (10) nor correctly explained on p11, and the cited Bishop (2000) is not in the Reference. [The cited paper could be Bishop et al. (2001) as listed in the Reference, but (10) is not seen or obviously seen there.] In particular, if “the vector  $V$  has gridded values of (10)” as explained in line 243, then  $V$  cannot be a linear projection operator (or any operator) as used subsequently in (10)-(22). As far as I can see,  $V$  should be a linear operator that projects the state vector  $X_e$  (in the full space  $R^L$ ) onto the subspace spanned by the complete set of vectors

that are composed of the gridded values of vector  $(u'/2^{1/2}, v'/2^{1/2}, c_p^{1/2}T'/T_r^{1/2})^T$ .

4. (19)-(20) on p14: “ $x_{iK}$ ” should be “ $x_{lK}$ ” for the last term in the first row of the matrix in (19) and (20). “ $\Psi^1$ ” is missed (perhaps due to a typo) on each side of the matrix in (19).

5. p 15 line 320: Change “Use (14), (15), (16), and (21)” to “Using (14)-(16) and (21)-(22)”. Clarify that “ $I_l$ ” defined in (21) has only one nonzero element, that is, 1 for the  $l$ -th diagonal element.

6. p19, line 421-423: I cannot see why “ET sensitivity is inversely proportional to the analysis error variance” as stated here. As far as I can see from (5)-(8), ET sensitivity is proportional (not inversely) to the analysis error variance.

7. p19-21 section 5:

The signal field (72h ahead on 200 hPa) in Fig. 3 is not explained but it appears to be the same as that in Fig. 4a, but the latter is produced by ETA from EXP\_Ag (rather than EXP\_U) on 500 hPa (rather than 200 hPa). This is very confusing.

The fields in Fig. 4 should be 72h ahead as in Fig. 2a (rather than 0h ahead “as in Fig. 2g” as stated in the caption of Fig. 4).

#### **Minor comments and suggestions:**

p9, line 190: Change “[8]” to (8)”.

P11: “ $P(v)$ ” is not defined. Change “ $P(v)$ ” to “ $P_e(v)$ ” in (11).

p13, line 288-291: Change “(14)” to “(15)” on line 288. Note that  $X_e(a)^+$  defined in (14) is not unique, and this may explain the “quite different behaviors”.

p13, line 295: Delete “for simplicity”.

p21, line 454: Delete “variation”.

Reviewed by Qin Xu